Preserving Bread Crumbs

Mary Baker
HP Labs

“All problems in computer science can be solved by another level of indirection.”

-- David Wheeler, 1927–2004

(World’s first PhD in CS, 1951)

But don’t forget the rest of the quotation:

“Except for the problem of too many layers of indirection.”

Maybe we should also add:

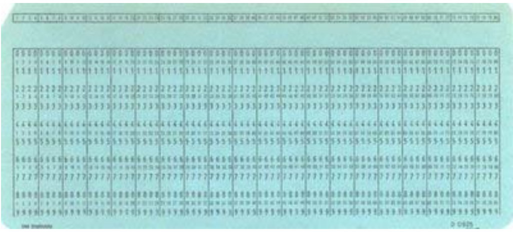
"Especially if it needs to keep working a long time from now."

Bread crumbs

- A level of indirection
  - To get here from there again in the future
- Problems
  - Birds eat them
  - They blow away
  - They decompose
  - We forget what they were supposed to mean
  - If there are too many, it’s a barrier

Digital preservation goals

- Digital assets stored now should remain
  - accessible
  - usable
  - undamaged



- for as long as desired – beyond the lifetime of
  - any particular storage system
  - any particular storage technology
- and at an *affordable cost*

SNIA 100 Year Archive Requirements Survey

68% of organizations had requirements > 100 years

83% of organizations had requirements > 50 years

Why it’s hard

- Large-scale disaster
- Human error
- Media faults

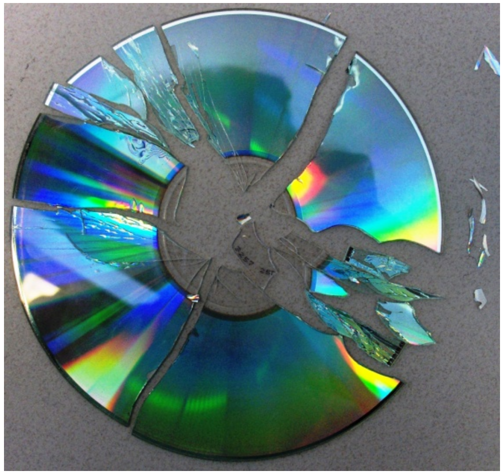
Long-term content suffers from more threats than short-term content

- Component faults
- Economic faults
- Attack
- Organizational faults

- Media/hardware obsolescence
- Software/format obsolescence
- Lost context/metadata

Why it's hard

- Large-scale disaster
- Human error
- Media faults



- Component faults
- Economic faults
- Attack
- Organizational faults

- Media/hardware obsolescence
- Software/format obsolescence
- Lost context/metadata

Why it’s hard

- Large-scale disaster
- Human error
- Media faults

- Component faults
- Economic faults
- Attack
- Organizational faults

- Media/hardware obsolescence
- Software/format obsolescence
- Lost context/metadata

Why it’s hard

- Large-scale disaster
- Human error
- Media faults

- Component faults
- Economic faults
- Attack
- Organizational faults

- Media/hardware obsolescence
- Software/format obsolescence
- Lost context/metadata

Even preserving the bits is hard

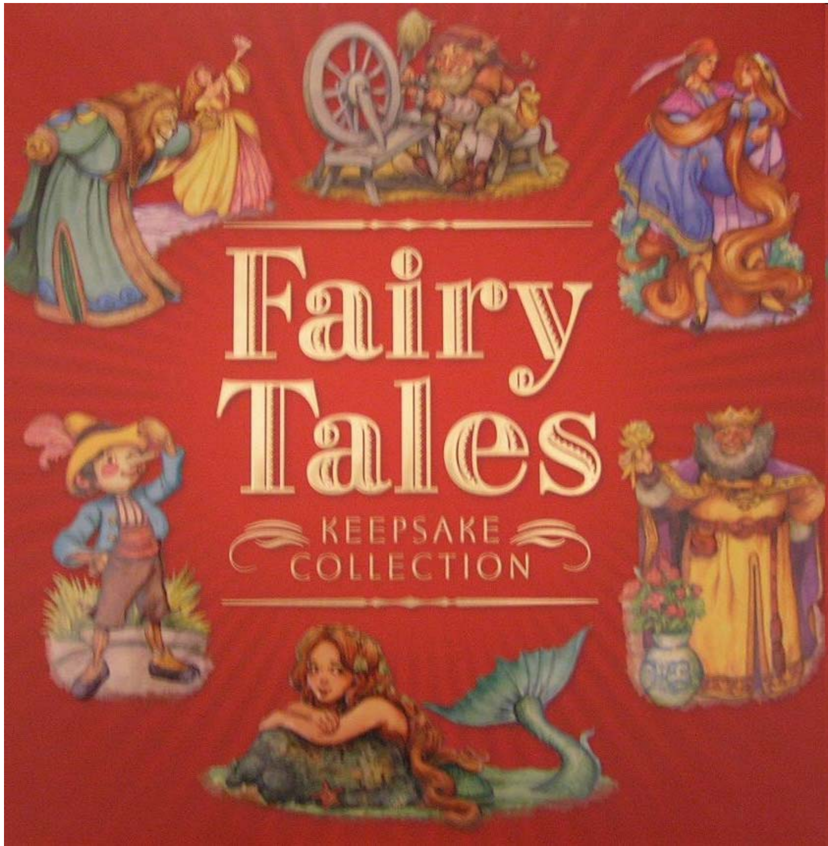
- Large scale & long time periods are a problem
- 1 petabyte, 50 years, 50% probability of no damage
  - Sounds reasonable, doesn’t it?
- That’s a bit half-life of 10^{17} years
  - A million times the age of the universe
  - Even improbable events will have an effect
- Now try to keep
  - The bits usable
  - The information reusable
  - The applications executable

Stuff we've tried

- Replication
- Error correcting codes
- Loss-tolerant formats
- Deduplication
- Virtual machines
- Canonical formats
- Self-describing formats
- Standardized data formats
- Format migration
- Preservation of ancient equipment
- Etc.

Some stories

- Virtual machines
- Replication with encryption
- DRM



The great U.S. springtime ritual

- Affects all segments of society
- Young and old
- Rich and poor
- Every culture, race, creed, religion
- Every profession

The great U.S. springtime ritual

- Affects all segments of society
- Young and old
- Rich and poor
- Every culture, race, creed, religion
- Every profession

Tax time!

How to do the taxes

- Need to
  - Get at past returns & even modify them
  - Reapply rules applicable at time created
- Used to save printed return & instructions
- But now we use Turbo Tax!
  - 15 years of Turbo Tax = 15 Turbo Taxes!
  - How can we run old copies?
  - Classic preservation problem
- Solution: VMs!

Mehul’s tax time story

- Installed VMware*, created virtual machine
  - Couldn’t load & install his version of Windows
  - OEM version of Vista – only one platform and only once
- How to get VMware to present new bios to Vista?
  - Various versions of VMware have various ways
  - Configuration parameters not intended for use
    - Undocumented
    - Change across VMware versions
- Machine image built for VMware x failed for VMware x+y
  - Need that particular version of VMware, or
  - Need to keep migrating machine image across VMware versions
  - New 92-page instructions on migrating to new VMwares!

* Not picking on VMware – it’s a problem for any virtual machine approach

Dirt under a different rug

- Just a different preservation problem
  - Migrate whole machine images
  - Instead of applications and content
- In addition
  - New VMwares might not run on old OS version
  - Many configurations of VMware for different platforms
  - Attacks: 10 years from now hypervisor converts photos to porn
- Similar problems handling (de)compression with VMs
- Still requires
  - Human thought
  - Planning
  - Process
  - Effort across time

Another rug story

- Replicate content widely for its survival
- Protect it with encryption
- Now you have other problems
  - Must preserve the keys and their usability
    - OceanStore: An Architecture for Global-Scale Persistent Storage, Kubiatowicz, et al., 2000.
  - Will need to
    - Preserve or migrate from old decryption algorithms/systems
      - pgp → gpg problem
    - Evolve/manage PKI/identity
      - Will my PKI last 50 years?
  - We may have to do this, but it doesn't come for free
    - Still requires human thought, planning, process, effort across time

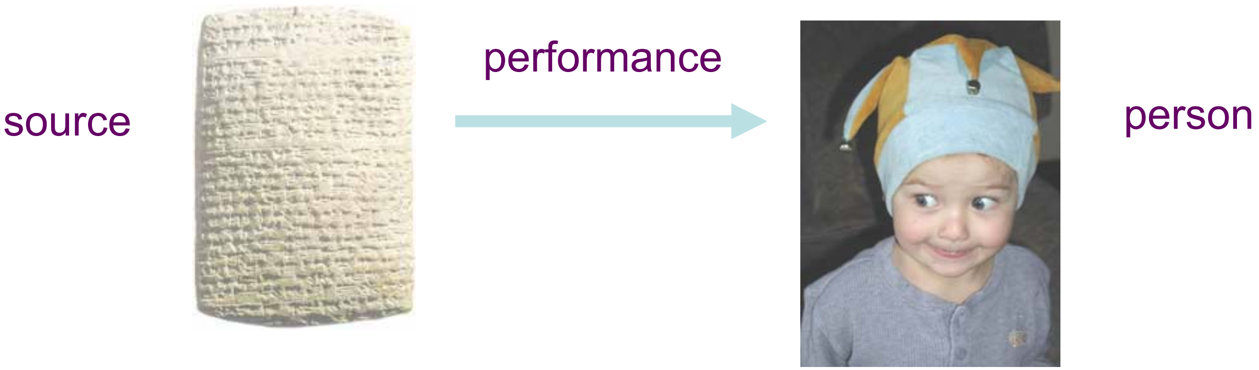
Nupur's story

- Took iPhone video of daughter's dance class
- Wanted to download to Windows XP PC
  - Just drag & drop!
  - After 40 minutes there's a 0 byte file
  - Try again – same thing
- DRM issue – locked down platform
  - Assures authenticity!
  - (Keeps her from stealing music & videos)
  - But this is her own content!
  - Had to go iPhone → Vista → XP
  - Give up and store on iPhones?

Fast forward 50 years

- Granddaughter wants that video
- Even if we still understand the format…
  - Wasn't a standard interface
  - Locked down for wrong reasons
  - There's too much in the way

Performance model



source

performance

person

- National Archives of Australia Performance Model
  - "A Performance Model and Process for Preserving Digital Records for Long-term Access", Andrew C Wilson of the National Archives of Australia, *IS&T Archiving Conf*., 2005.
- Less technology for performance → easier to preserve
  - Don't have to preserve that extra technology
  - Avoid indirections and barriers
- Requires human knowledge to persist
  - Widely understood languages, standards help

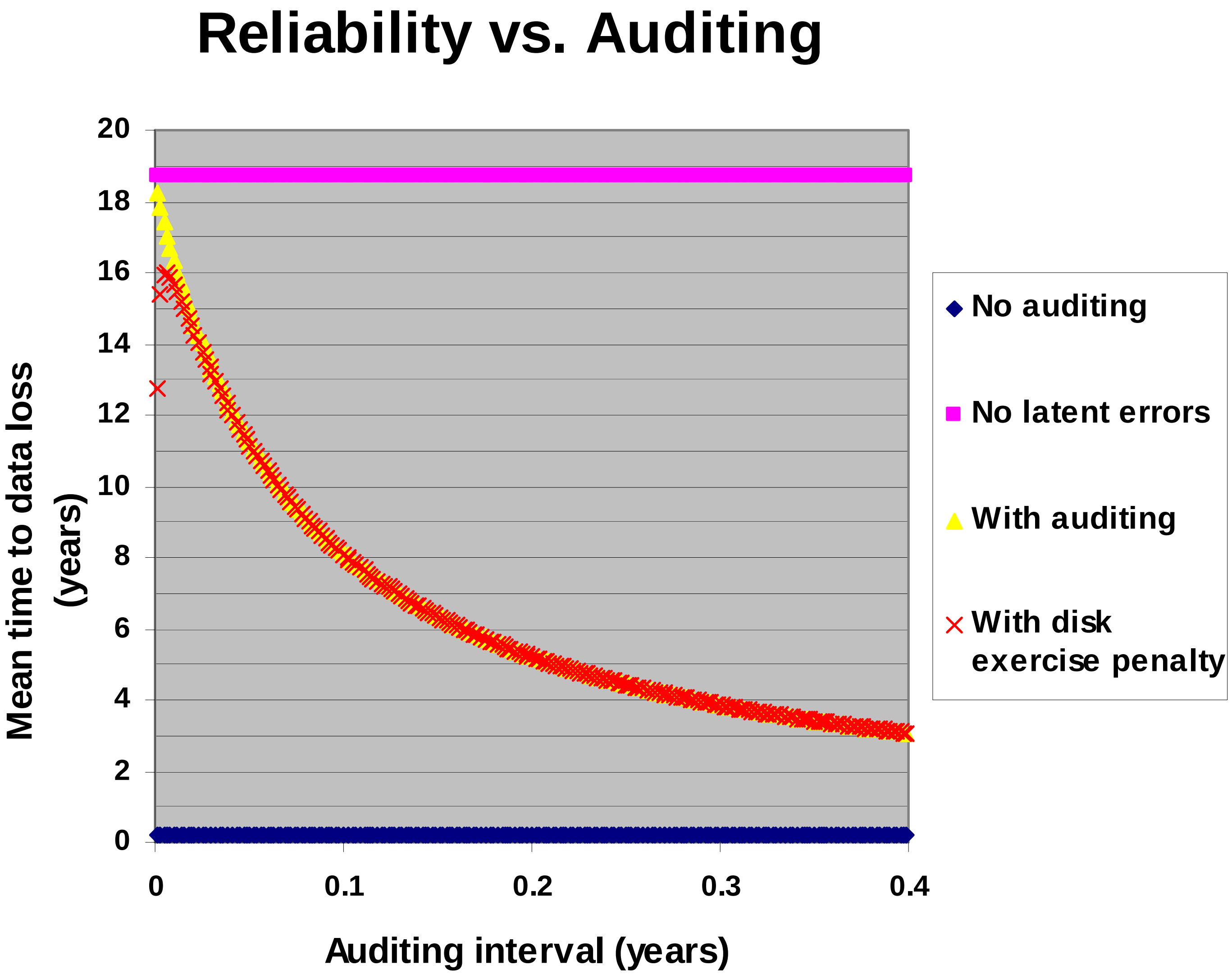
Model explains, but doesn’t solve

- Decide on *essence* to perform
  - What do you need to capture and replay?
- Ensure you continue to produce performance of essence
- Often can’t reproduce original performance exactly
  - Examples: video games, old books
  - Cost: how much of essence can you capture at reasonable cost?
  - Future recipients might not share creators’ view on essence/value
    - Don’t know what will be essential in the future
- Essence to preserve can depend on the
  - Domain
  - Eventual purpose
  - Funding

Key ideas / best practices

- Replicate content
  - If one copy is damaged, can repair from another
  - It's not enough to make a single "super reliable" copy
- Avoid correlated failures
  - Not just geographic, but administrative, platform, etc…
  - Heterogeneity helps avoid correlations
  - Must balance this with cost and administrative hassle
- Find & fix (if possible) latent faults before damage grows
  - Some faults don't announce themselves
  - Latent faults can occur at all layers, physical and logical
  - Must look for silent damage/problems proactively: "audit"
  - Includes auditing for pending obsolescence, etc.
  - This means the content must be accessible!
- Use widely understood standards
  - Help customers avoid metadata and format traps

Example: audited, replicated archive



Baker et al., A Fresh Look at the Reliability of Long-term Digital Storage." *EuroSys 2006.*

Other techniques

- Deduplication
  - Reduces cost of geographically independent replicas
  - Don’t dedupe across sites
  - Don’t make audit too expensive
  - Take care with metadata and system migration
- Self-describing formats (SIRF)
  - Good to embed human understanding
  - The bread crumbs are the registries
  - Still requires attention over time
- Reformat on demand (LOCKSS)
  - Avoids cost of reformatting stuff we don’t access
  - Bread crumbs are the conversion functions
- Format migration (everywhere)
  - It’s a way of formalizing continued attention
  - Error-prone, may end up with poor concept of original
  - Asserting authenticity, chain of custody, integrity won’t come for free
- Swiss Fort Knox time capsule (Planets)
  - This really is a barrier

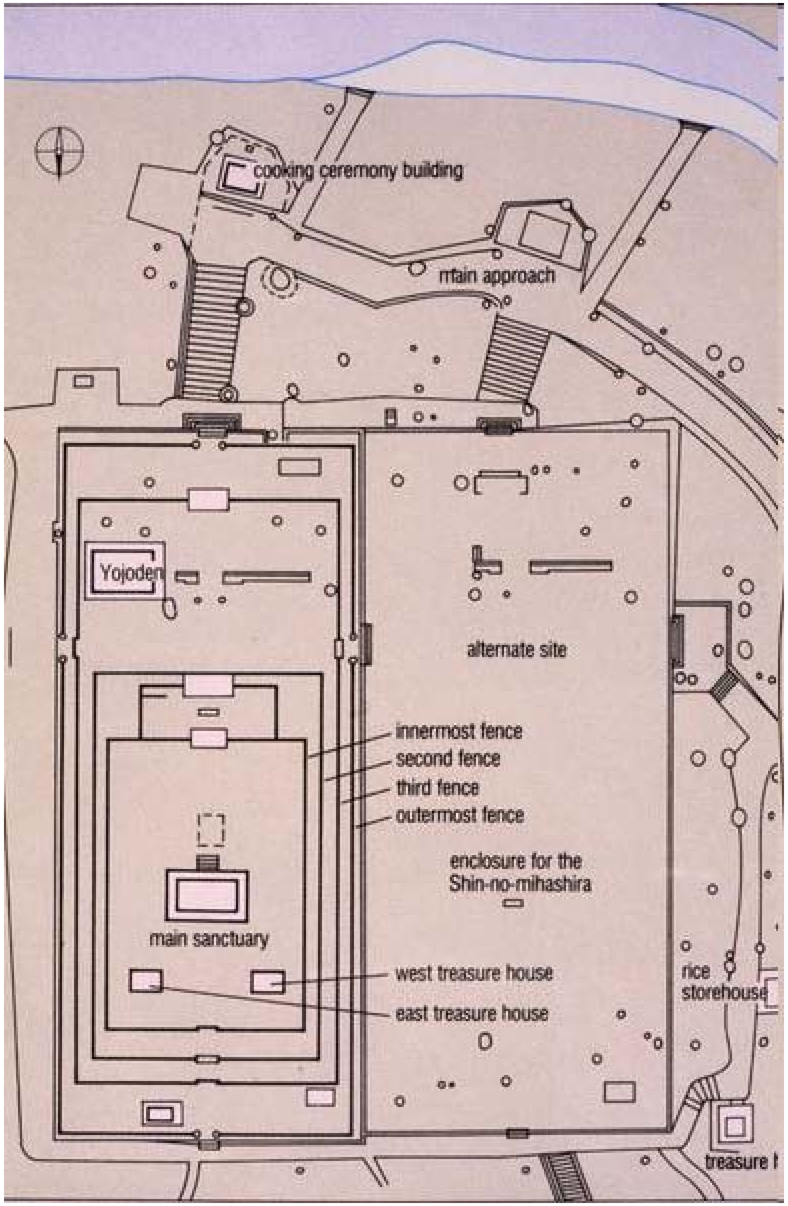
What have we learned?

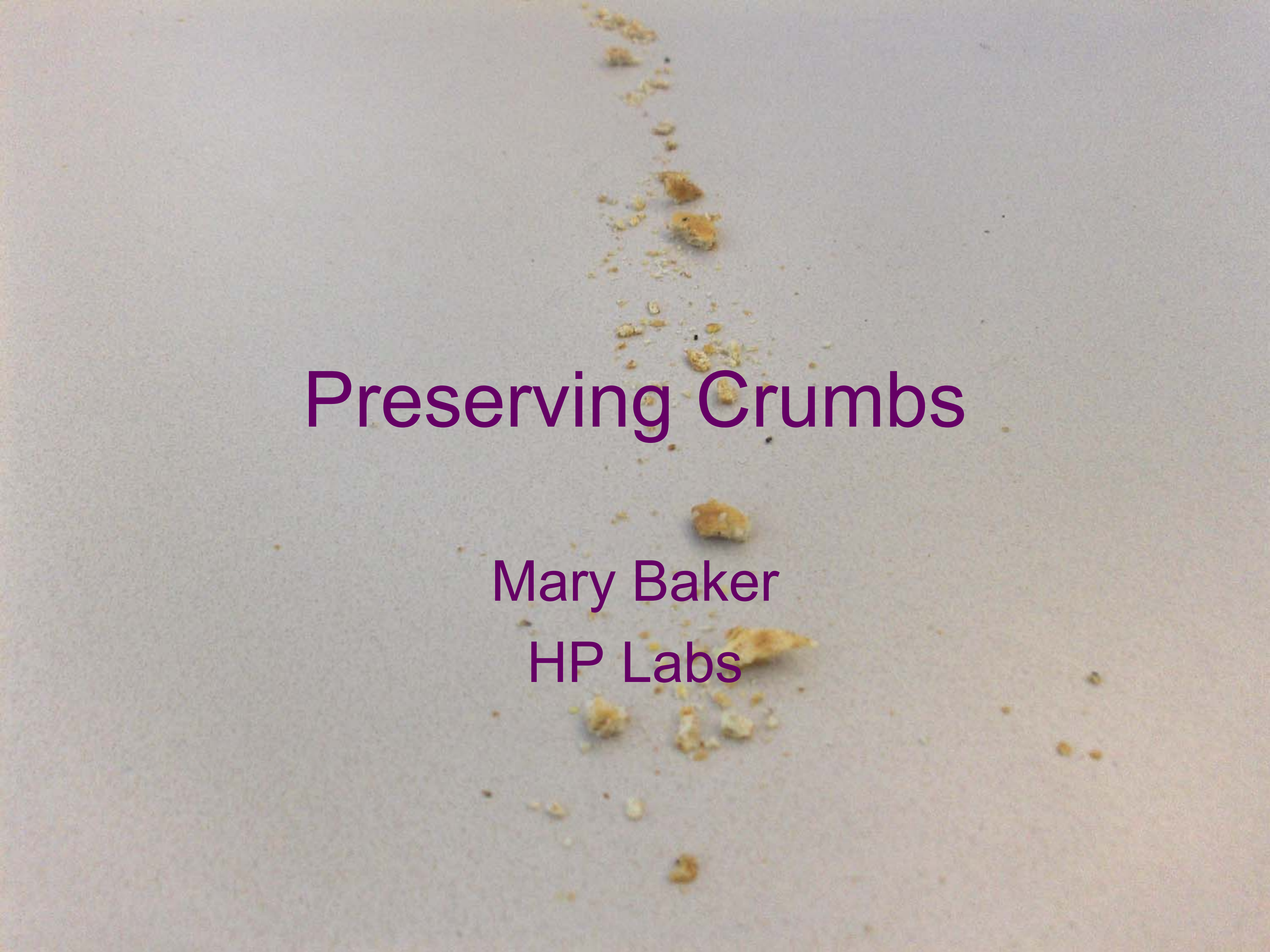
- Another layer of software is another layer of software
- Sweeping dirt under the rug doesn’t ensure victory
- Interfaces should provide access, not put up barriers
- Don’t protect what you don’t need to protect
- It helps if a lot of people know about it
- Check for trouble before it’s too late
- Every indirection increases vulnerability
- No way to avoid human attention at some point

What can we do?

- Keep checking
  - The content is still okay
  - The processes still work
- Avoid secrets, proprietary formats/APIs
- Evolvable, transparent processes
  - Especially for secrets and proprietary stuff
  - We don’t know what will change, but change it will
- Understand that human attention is essential
  - Figure out where to focus it
  - Maybe reduce frequency/cost of that attention
- But we can’t leave it alone!
- Keep rebuilding the Sanctuary of Ise

Sanctuary of Ise -- Naiku (inner shrine)



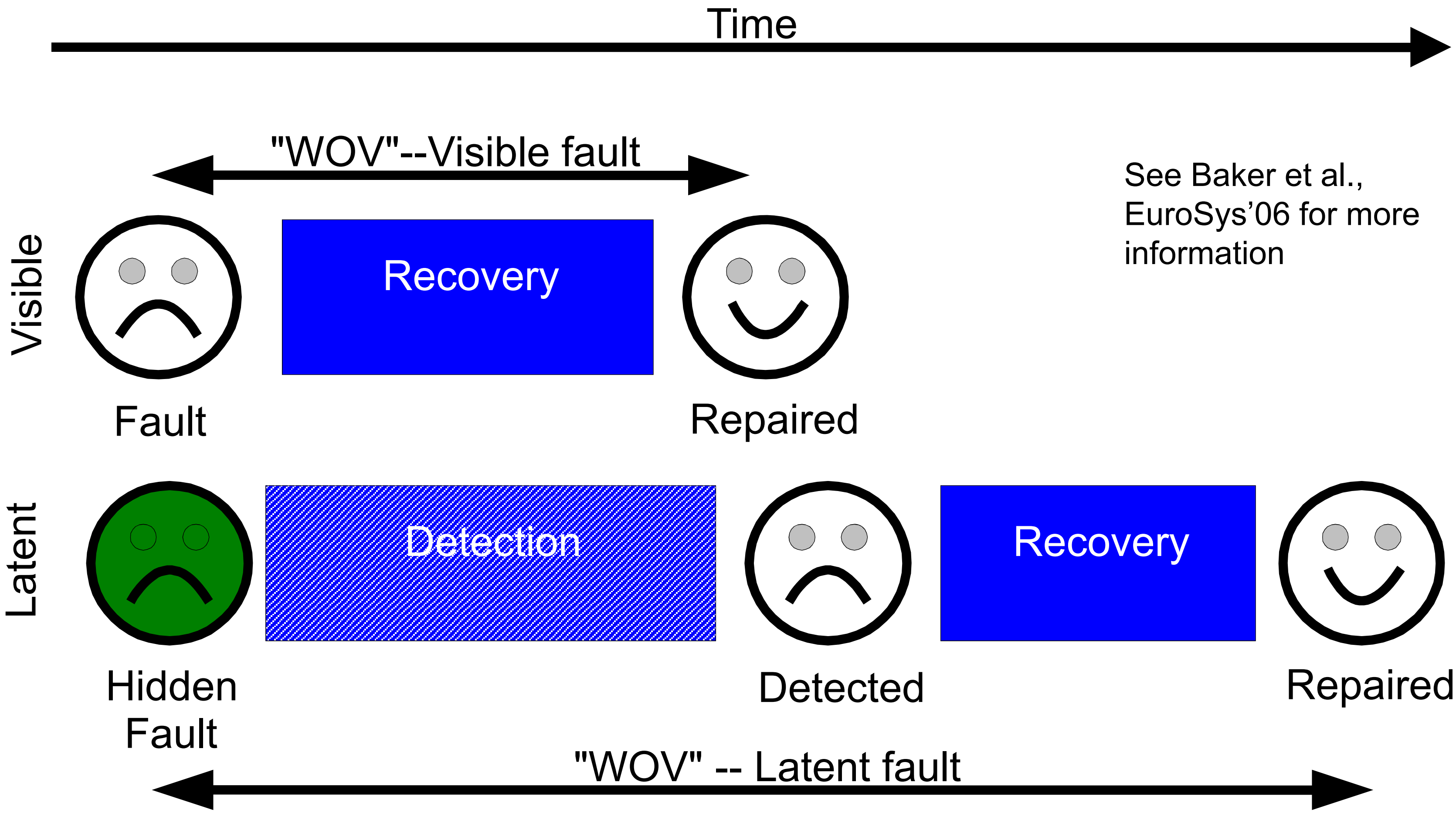


Preserving Crumbs

Mary Baker
HP Labs

Window of vulnerability

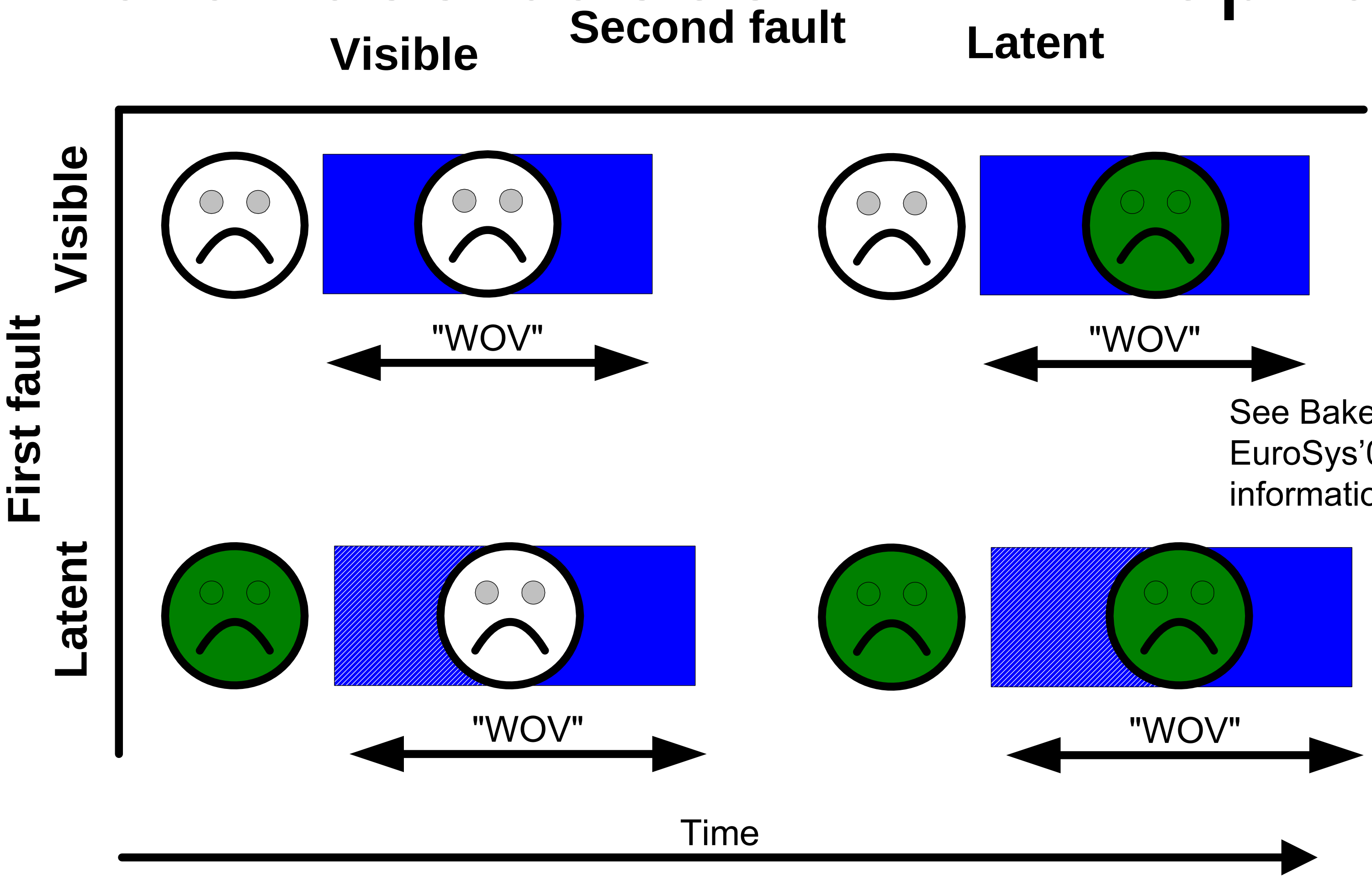
Temporal overlap of faults



See Baker et al., EuroSys'06 for more information

- Want detection time to be small

Data loss cases with 2 replicas

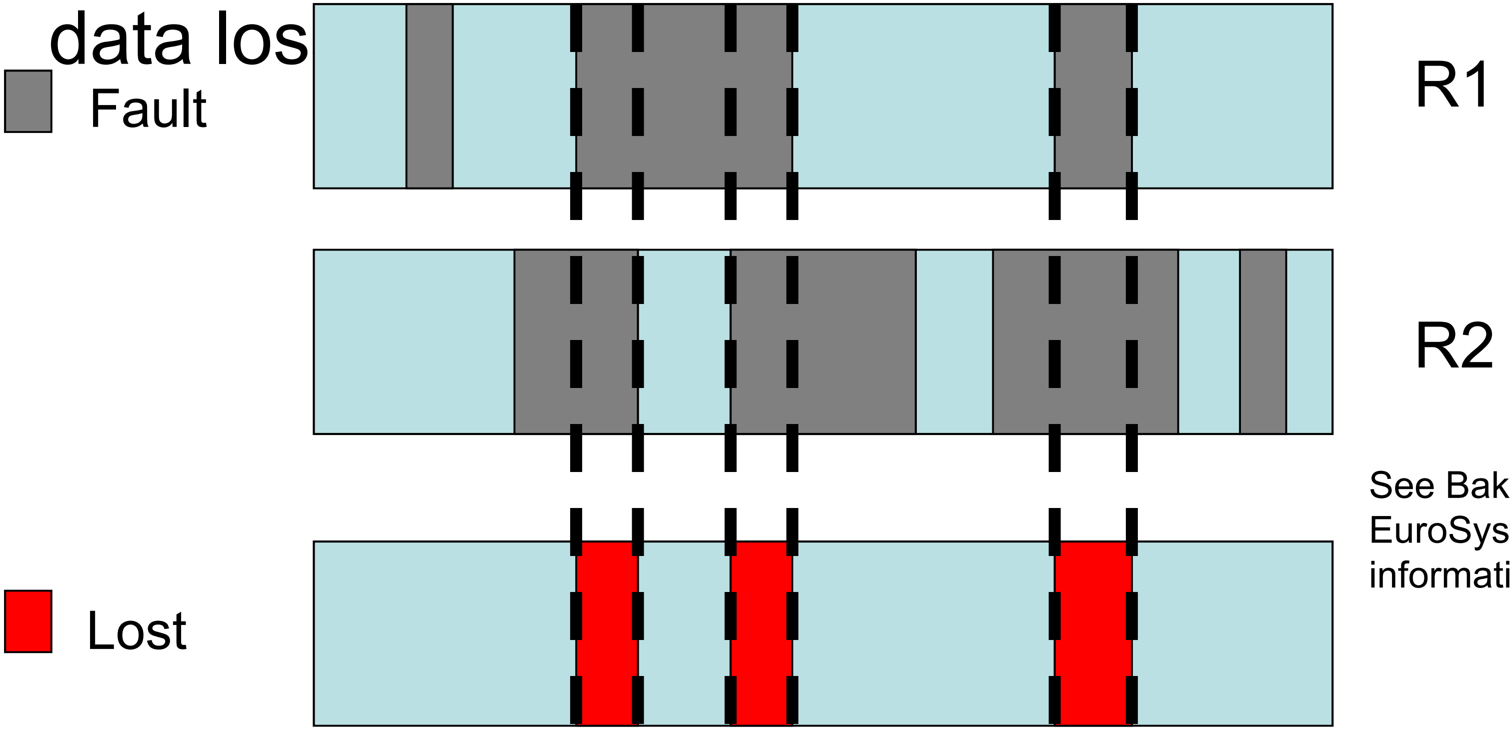


See Baker et al., EuroSys'06 for more information

- Overall probability = sum of each case

Spatial overlap of faults

- Temporal overlap alone overstates likelihood of data los



See Baker et al., EuroSys’06 for more information

- Faults may be bits, sectors, files, disks, arrays, etc.
- If any two faults overlap, data is lost
- The smaller the faults, the less likelihood of overlap